

White Paper

Grant Number: HD-51863-14

Project Title: Humanities CORE

Project Directors: Kathleen Fitzpatrick, Modern Language Association
Barbara Rockenbach, Columbia University Libraries/CDRS

Grantee Institution: Modern Language Association of America

Report Date: 27 August 2015

Name of Project: *Humanities Commons Open Repository Exchange (HumCORE)*

Institution: Modern Language Association of America

Project Directors: Kathleen Fitzpatrick, Modern Language Association; Barbara Rockenbach, Columbia University Libraries/CDRS

Award Amount: \$60,000

To Support: The development of software to connect the *Commons-in-a-Box* (CBOX) social network platform (which is the basis of *MLA Commons*) to a Fedora-based institutional repository system. This combined system would be called *Humanities Commons*, a social network and repository system that would be made available for use by other scholarly societies.

Project Web Site: commons.mla.org/deposit

Keywords: repository, collaboration, grey literature, social knowledge environments

Grant/Period: NEH Digital Humanities Start-Up Grant, 05/01/14–05/31/15

White Paper to the NEH Office of Digital Humanities *Humanities Commons Open Repository Exchange*

White paper authors: Nicky Agate and Mark Newton

ABSTRACT

The Modern Language Association (MLA) and the Columbia University Libraries / Information Services' Center for Digital Research and Scholarship (CDRS) have been working together on a prototyped user interface that connects a library-quality repository system (*Humanities Commons Open Repository Exchange*, or *CORE*, modeled on Columbia's *Academic Commons*) with *MLA Commons*, an academic social networking platform that enables MLA members to communicate, collaborate, and share their work with one another. The *CORE* interface allows *Commons* members to upload, share, discover, retrieve, and archive digital work and other objects in an existing collaborative network. This white paper describes the team's progress and setbacks over the grant period.

INTRODUCTION

In early May 2015, the Modern Language Association released the beta version of *Commons Open Repository Exchange (CORE)*, a new social repository for members of the scholarly society. *Commons Open Repository Exchange* is a prototyped user interface that combines a library-quality repository for sharing, discovering, retrieving, and archiving digital work (modeled on Columbia University's *Academic Commons*) with the social networking capabilities of *MLA Commons*, the Modern Language Association's in-development network for scholarly communication, collaboration, and publishing.

CORE is the result of a yearlong collaboration between the MLA and the Center for Digital Research and Scholarship at Columbia University, whose expertise in repository development and management was crucial to its success. Funded by a \$60,000 Digital Humanities Start-Up Grant from the NEH, the collaboration resulted in the full specification for and a working prototype of a *WordPress* plug-in that permits the repository, whose design is based on that of Columbia's *Academic Commons*, and *MLA Commons* to interact seamlessly. This unified interface provides access to collective functions including file uploading, metadata creation, discovery, and retrieval, all within the front end of the *Commons*. The interface front end is managed by a *WordPress/BuddyPress* plug-in connecting CBOX's file management system to the Fedora/Solr repository. The result is robust content indexing, discovery, display management, and long-term archiving, and, since this solution is open source, other groups using a Fedora/Solr stack for asset management and discovery who wish to integrate a repository with an active collaboration platform built in CBOX will be able to use it too. In essence, then, *CORE* pairs the user-friendly, commercial-grade, open-source community-building software of the *Commons* with library-developed and -tested technologies that enable the search, discovery, access, and preservation of scholarly material and serves as a gateway to browse the faceted contents of shared materials, to discover related contributions, and to facilitate their possible reuse.

Like other repositories, *CORE* facilitates open-access distribution, discussion, and citation of the many products of humanities research, including pre- or postprints, conference presentations, data sets, and learning objects such as syllabi and slide decks. What makes *CORE* stand out, however, is its social facet, the fact that it is not an independent entity but an integrated part of *MLA Commons*, the very environment in which MLA members already connect with one another. The MLA boasts almost one hundred and fifty forums dedicated to disciplinary and interdisciplinary subfields of language and literature. Members choose five of these upon joining, with which they become officially affiliated; they may, however, join the *Commons* groups for any others in which they are interested. Involvement in these forums gives MLA members an immediate audience of interested peers for their work, and ensuring a connection between *CORE* and a member's forum memberships was instrumental in the development of this prototype. The integration between the association's forums, *MLA Commons*, and *CORE* aims to transform the repository from a static to a dynamic and durable platform for open scholarly communication.

MLA COMMONS

MLA Commons is a Web-based scholarly communication, networking, and publishing platform built on the open-source *WordPress/BuddyPress* software Commons-in-a-Box (CBOX) and launched with the generous help of the Andrew W. Mellon Foundation in January 2013. The platform is an example of what Nancy Fjällbrant calls "the invisible college" (2), an informal scholarly network of group and single-author blogs and Web sites. It features discussion areas for members of the MLA's many official disciplinary forums (covering areas as far afield as

global and Arab American literature, memory studies, history and theory of composition, and Marxism, literature, and society), as well as member-created special-interest groups. And it offers a robust portfolio system that enables its 5,200 members to cultivate a professional online presence in the heart of their disciplinary society.

The more traditional MLA publications, including the association's journals and book series, are exploring the potential presented by the *Commons* both for new means of developing in-process publications and new forms of outreach and engagement after publication. Among the numerous projects and conversations already being hosted by the platform are the MLA's first all-online, evolving collection of essays, *Literary Studies in the Digital Age* (dlsanthology.commonsm्ला.org); *Profession*, a journal focused on trends in modern language and literature scholarship and pedagogy (profession.commonsm्ला.org); *Reforming Graduate Education*, an open discussion of the 2014 Mellon-funded task force report on that subject (gradreform.commonsm्ला.org); and three book projects in development, *Digital Pedagogy in the Humanities: Concepts, Models, and Experiments* (digitalpedagogy.commonsm्ला.org), *Teaching Anglophone South Asian Women's Writing* (teachingsouthasia.commonsm्ला.org), and *Teaching Modernist Women's Writing in English* (modwomen.commonsm्ला.org). MLA members are also prototyping online open-access peer-reviewed journals to be published on the platform.

THE ORIGINS OF CORE

Since the first scholarly society published the first academic journal, the academic article has reigned as the primary form of official scholarly communication (Fjällbrant 3); it seems appropriate, then, that the MLA, the largest scholarly society in the world, would now create the first disciplinary repository for scholarly communication in all its forms.

Traditional publishing, in a knowledge environment that is not only increasingly digital and increasingly networked but also increasingly open, is failing today's scholars in many ways. Whether because of the increasing costs of journals and cuts in library budgets that result in restricted access to scholarship, the eighteen-month time lag between the submission of a manuscript and its publication, the potential intransigence of peer reviewers vis-à-vis new ideas, or the inability of traditional publishing to adapt to nonprint or hybrid forms of publication, scholars in the humanities are demanding new kinds of access, both to resources and to one another. Many academics today are doing scholarly work in ways that *exceed* the capabilities of established forms like the book and the journal article: they are posting ideas to blogs, tweeting at conferences, cleaning data sets for text mining and data visualization, and creating digital learning objects for the classroom. These, too, are forms of scholarly communication, albeit "emergent" ones ("Research Blogs"), and as such deserve to be both valued as professional contributions and made available—to be made public, the original sense of "publish," as Karin Wulf recently reminded us—in a social knowledge environment such as the *Commons*. To this end, the Modern Language Association has partnered with EZID, the University of California Curation Center's provider of unique, long-term identifiers, to ensure that every object deposited

is granted a DOI, thereby establishing a permanent record of originality and authorship for the depositor.

CORE BETA

Because the level of skill and comfort with technology varies widely from member to member, *CORE*'s interface was designed to be simple, user-friendly, and as familiar as possible. To the user, then, *CORE* appears as a fully integrated element of *MLA Commons*, a platform with which members, at least, are already acquainted. *CORE*'s presence in the main *MLA Commons* menu gives it equal footing alongside the platform's frequently visited pages: the group, member, and site directories; the activity feed; information about *Commons* publications; and the Help page (see fig. 1).

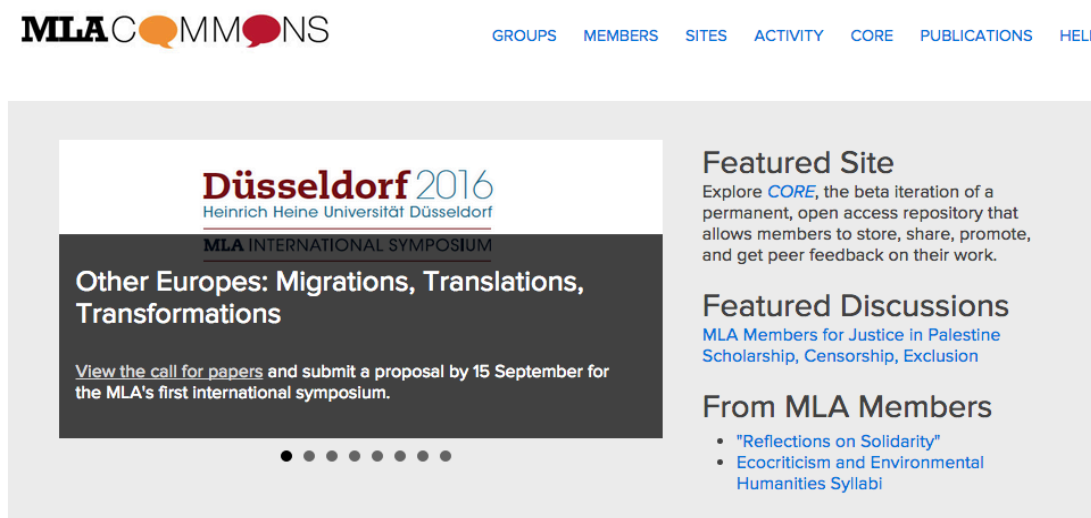


Fig. 1. *CORE* in the top-level navigation of *MLA Commons*

While anyone can view and download objects deposited with *CORE*, only *MLA Commons* members can upload their work to the repository. Scholarly work here is broadly defined, and the interface will accept anything from maps to syllabi to visual art to pre- and postprints of published articles (see fig. 2).

New CORE Deposit

Select the file you wish to upload and deposit. *

Select File

Cancel

Title

Item Type

Select an item type ▲

- Code or software
- Conference paper
- Conference proceeding
- Course material or learning objects
- Data set
- Dissertation
- Documentary

Fig. 2. Items that can be uploaded to *CORE* run the gamut of forms of scholarly communication

The user can accord item-level permissions to each object deposited: the assignment of any one of the available Creative Commons licenses and the inclusion—or not—of the object deposited in an MLA forum's library on the *Commons* (see fig. 3). As activity on forum pages triggers (preference-dependent) e-mail notifications to members of that forum, the act of depositing an item with a forum means the author has an immediate and self-nominated community of interest to serve as an initial audience for his or her work. And since the deposit can be followed up by a forum post asking for feedback or giving context to the work, the author can draw additional attention to something while inviting open discussion within a *Commons* group. Finally, the availability of view and download metrics enables all users to evaluate the success of such efforts.

Interdisciplinary Approaches to Culture and Society Group Admins

Discussion Group (Public) active 37 minutes ago

Leave Group

Deposit an Item

Order By: Newest Deposits

The Interstitial Body and Moral Formation: Third-Culture Displacement and Subject Formation in Charles Kingsley's The Water-Babies

Author(s): [Hannah Swamidoss \(see profile\)](#)

Date: 2013

Forum(s): [Interdisciplinary Approaches to Culture and Society](#), [Late-Nineteenth- and Early-Twentieth-Century English Literature](#), [Literature and Religion](#), [Literature and Science](#), [Popular Culture](#), [The Victorian Period](#)

Subject(s): [British literature](#), [English literature](#), [Literature and science](#), [Religion](#)

Permanent URL: <http://dx.doi.org/10.17613/M6QC73>

Fig. 3. Users can opt to include their items in the deposit library of up to five MLA forums

Additional discoverability is enabled by the openness of the repository and the *BuddyPress* activity stream functionality that has been tightly woven into the fabric of *CORE*. Objects deposited in the repository are discoverable not only on the *CORE* page and on the home pages of various MLA groups but also through the global Activity menu on the *Commons*. User profiles, like groups, have a new deposits facet that enables members to view and download the work of others with whom they might be interested in connecting or collaborating, and since these deposits can be accessed from within the author's profile, they are discoverable by users of the platform who might not otherwise visit the *CORE* section of the site.

PLATFORM

The *CORE* user interface connects the *WordPress*-based CBOX platform in use at *MLA Commons* (commons.mla.org) and a Fedora/Solr repository platform. This interface between the *Commons* and *CORE* provides access to collective functions including file uploading, metadata creation, discovery, and retrieval, all within the front end of the *Commons*. The addition of a library-quality repository forms a key component of the ongoing development of *MLA Commons*, in which the MLA seeks to open the collaboration and communication functions of

the *Commons* to other scholarly societies through the creation of a federated *Humanities Commons*.

To develop the *CORE* prototype, the MLA partnered with the CDRS team, whose expertise in repository development and management was crucial to its success. The result of our collaboration is a full specification for and a working prototype of a plug-in that permits the two systems, CBOX and Fedora/Solr, to interact seamlessly. The interface front end is managed via a *WordPress/BuddyPress* plug-in connecting CBOX's file management system to the Fedora/Solr repository. The result is robust content indexing, discovery, display management, and long-term archiving; this solution can be applied by other groups using a Fedora/Solr stack for asset management and discovery who wish to integrate the repository with an active collaboration platform built in CBOX. In addition to providing an interface for content addition and description between the *Commons* and *CORE*, the plug-in serves as a gateway to browse the faceted contents of shared materials, to discover related contributions, and to facilitate their possible reuse.

The target production environments for *CORE* and the pilot's reference system, *Academic Commons*, while sharing components in the underlying technology, are significantly different. The former relies on unmediated user interaction initiated with a *WordPress*-based site; the latter relies on separate platform components for the presentation, submission, and management of repository content. Both of them, however, rely on *Fedora Commons* (currently, version 3) for the archival and preservation layer.

The library's attention has already turned to planning for a significant upgrade of its Fedora-based projects to version 4 of *Fedora Commons*. The question to the *CORE* pilot project team: whether and how to plan for the commensurate, eventual upgrade of the *CORE* repository as well. After consultation with digital content architects in the Libraries Digital Program Division, the decision was made to adopt a more flexible content model for resources and metadata, thus smoothing the path to eventual platform migration.

LESSONS LEARNED

The collaborative process behind the creation of *CORE* has been an overwhelmingly positive one, but we have not been without our challenges. Among these is one that is not unique to our repository: how do we encourage humanities scholars to deposit their work when they are used to exposing finished *products*—the journal article, the book—rather than process or works in progress?

The work of academic libraries to build open repository programs for their local constituents is as long as it is documented. Build it, but will they come? Staff it, but at what levels and with which skill sets? Make it easy to use, but leverage the moment to change scholar practices in the negotiation and retention of publication and distribution rights. Administrators of such programs look to the bottom line: demonstrable evidence of engagement and value perception with scholars *as measured in total number of deposits*. While libraries have done significant work in making the research output of their faculty and students available to

university stakeholders, the institutional repository often poses discoverability and access issues when it comes to the broader community. A scholar's work may be publicly and openly available, but—and especially if that scholar is more junior or the institution less renowned—availability does not equal visibility. Growth rates for repositories, where contribution is voluntary and beyond the reach of a faculty resolution for open access, remain anemic. As Nathan Hall points out in reference to institutional repositories, “many university faculty have rather conservative attitudes towards scholarly communication” (339). Hall's research for his 2015 study involved an informal survey of faculty members that revealed “many participants were unaware of institutional repositories and understood open access only in terms of hybrid journals” (341).

Faculty members who publish place a value on contributing content to open repositories that is determined to a significant extent by disciplinary practice (as opposed to institutional imperative). The *MLA Commons–CORE* connection is an important bellwether for the potential of disciplinary repositories to increase distribution of content when connected to hubs for collaboration.

From the perspective of the stakeholders at the Columbia Libraries, the *CORE* pilot has been a marked success. The project suggests that partnership with allied groups such as scholarly societies can be effective when long-term goals—such as the promotion of scholarly collaboration and access to published content—are well aligned between the groups. Columbia's role in the *CORE* pilot was largely *consultative* (with respect to experience in building repository and engaging users) and *connective* (bridging technology teams to connect the MLA development work with Columbia repository components), and during the course of the project we learned a lot about technology and platform, connections with scholars, and refinements to the vision of the road ahead in extending beyond the initial collaboration.

It is too early in the life of *CORE* to fully understand the full range of benefits—and, perhaps, challenges—the plug-in presents to MLA members. At the time of this white paper's writing, the deposits numbered twenty-eight articles, five conference papers, five book chapters, four essays, four syllabi, three dissertations, two theses, two bibliographies, and a book. Every item has been shared with at least one MLA forum, and most have been shared with more than one group. Some items deposited in the initial days of the prototype's release have been viewed over three hundred times and downloaded by more than thirty people.

When it comes to discoverability, the advantages seem obvious. Members who have recently given a well-received talk at a conference can share their work with other scholars in their field who, for whatever reason (often financial), could not attend. Graduate students for whom the timeframes of traditional publication prove too long to be beneficial on the job market will benefit from the opportunity to bring their work to the attention of search committees and editors. Independent scholars and researchers, who have no means of affordably accessing a wide range of traditionally published work and whose status often negatively affects their chances of publication in major journals, can participate fully in the conversations of their field.

FUTURE DEVELOPMENT

In the next steps for the *CORE*, we will work to reconcile the institutionally bound vision with the disciplinary view in developing openly accessible scholarship. *CORE* provides an excellent launchpad for further investigation and development of workflows that begin with community engagement and result in content available through institutionally supported repositories. Eliminating redundancy and blocks to participation in open repository projects may be significantly improved through partnerships with society-sponsored digital communities. Our attention therefore turns to the careful business of crafting useful, efficient, and intuitive means of participating in such integrated systems.

Moving forward, we will need to address the issue of linking iterative versions and different components of scholarship together. *CORE* cannot currently display the connection between a data set and a research paper based on its findings, for example, or a conference presentation that served as an early draft for a traditional publication. Of course, the direction we take here will be part of the larger conversation around the existence and establishment of a version of record—if such a record exists (Wulf).

We will also have to tackle the question of subject fields, which in the beta release have been limited for the sake of simplicity. Do we allow user-defined fields, and, if so, will these take away from the value of the platform's search functionality? Can we create an evolving subject taxonomy to help users categorize their content, thereby standardizing metadata and reducing the need for later deduplication? And, given that the MLA does not have a librarian or other metadata expert on staff, how can we ensure that members are entering the metadata necessary for other users to be able to correctly cite their material?

Of course, feedback from MLA members over the remainder of the beta testing period will be the foundation of future development plans. One of our principal areas of focus, however, will be the development of a shared repository for a multisociety *Commons* network whose creation will be funded by a grant from the Andrew W. Mellon Foundation. Susan Schreibman, Ray Siemens, and John Unsworth have argued that the digital humanities have an “interdisciplinary core” (xxiv); we would suggest that increasingly it is not just the digital humanities but the humanities at large that has interdisciplinarity at its center. A federated *CORE* will harbor the potential for connections across several *Commons* instances and allow the MLA, with its relatively mature *Commons* model, to host such instances for organizations that cannot host their own.

By opening both the network and the repository to other scholarly societies, the MLA will answer our members' call for increased interdisciplinarity. A federated *CORE* would go some way toward opening access and enhancing discoverability of research across disciplines, enabling scholars to find connections between their work and broad issues within the humanities.

Bibliography

- Fjällbrant, Nancy. "Scholarly Communication—Historical Development and New Possibilities." *Proceedings of the IATUL Conferences*. Paper 5. 1997. *1997 IATUL Proceedings*. Purdue U, n.d. Web. 9 May 2015.
- Hall, Nathan. "The Information Literacy Gap in Scholarly Communication." *Virginia Tech Digital Repository*. VTechWorks, Mar. 2015. Web. 10 May 2015.
- "Research Blogs." *Types of Digital Work Wiki*. MLA, 11 Nov. 2011. Web. 10 May 2015.
- Schreibman, Susan, Ray Siemens, and John Unsworth. "The Digital Humanities and Humanities Computing: An Introduction." *A Companion to Digital Humanities*. Ed. Schreibman, Siemens, and Unsworth. Oxford: Blackwell, 2004. *Alliance of Digital Humanities Organizations*. Web. 26 Aug. 2015.
- Wulf, Karin. "Version Control; or, What Does It Mean to 'Publish'?" *The Scholarly Kitchen*. The Scholarly Kitchen, 16 Apr. 2015. Web. 8 May 2015.